

Ethics and fraud

The trajectory of the Hwang scandal highlights the shortness of the path between unethical behaviour and outright misconduct.

The fall of Woo Suk Hwang represents perhaps the highest-profile case in the sorry history of research misconduct. The sheer Shakespearian drama of the Korean cell biologist's eclipse, surrounded by fawning courtiers and plotting groups of acolytes and enemies and in full view of the television cameras, has left few researchers of any discipline, anywhere in the world, unaware of its circumstances.

But what really makes the Hwang case special is the importance of the impugned results (W. S. Hwang *et al. Science* **303**, 1669–1674; 2004 and **308**, 1777–1783; 2005). The claimed cloning of human embryos placed Hwang at the forefront of stem-cell research, perhaps the most acclaimed and contentious sphere of modern science. It is also the first major misconduct case to occur during the modern era of carefully patented biology, in which scientists may aspire not just to fame, but to fortune as well.

All of this raises some general questions about ethics and fraud that researchers had perhaps hoped were put to bed a long time ago. Many of these questions were last publicly addressed in the long and painful aftermath of the Baltimore case, in which a junior researcher, Thereza Imanishi-Kari, was accused of fraud in the laboratory of one of the United States' leading microbiologists.

David Baltimore was eventually vindicated and is today president of the California Institute of Technology. But when the allegations were made, the National Institutes of Health opened an Office of Scientific Integrity, which was later downgraded but continues to support fraud investigations at US universities while seeking to get academics to teach their students about ethics and misconduct.

"The US system for dealing with fraud, imperfect as it may be, is still more advanced than that of many other nations."

This system, imperfect as it may be, is still more advanced than that of many other nations. Elsewhere in the world, cases of fraud have highlighted considerable shortcomings in the mechanisms for responding to misconduct.

Another celebrated fraud case involved Jan Hendrik Schön, a physicist at Bell Laboratories in Murray Hill, New Jersey, who was found guilty in 2002 of fabricating results in a massive and influential string of papers on superconductivity. However, the general view of the Schön case, outside the discipline directly involved at least, has been that no innocent civilians were hurt, and the fraud was unearthed and duly punished. There is almost a nagging sympathy for how such a smart young man could be so stupid.

The scope of the Hwang scandal tends to leave these and other recent misconduct cases in the shade, however. Although the full facts of the case remain unknown, its multidimensional outline is already clear. Barely a stone was left unturned in his lab's attempts to secure a positive result. Egg donors were energetically sought from

every conceivable source — including from junior researchers inside the laboratory itself, where allegations of coercion have been made. In the end, a result was claimed and the paper published in the complete absence of genuine scientific evidence to support it.

This will surely leave some people asking: if this single cell in the body of science was so malignant, how fares the rest? This is an awkward question and one that many in the community will seek to evade by referring to the fact that it happened in South Korea and couldn't happen here (wherever 'here' happens to be). And it is true that standards of oversight at many laboratories would, at least in theory, make this kind of scenario improbable. However, research — not least in the dynamic and fiercely competitive field of reproductive biology — is increasingly conducted through international collaborations. These often involve working with colleagues in countries, including China and South Korea, where mechanisms for supervising ethics and investigating misconduct are at relatively early stages of development.

"This will surely leave people asking: if this single cell in the body of science was so malignant, how fares the rest?"

It therefore falls to scientists who take research ethics seriously to confront the question head on, and to determine what steps should be taken to redouble our efforts to secure standards in both the ethics and conduct of research.

Hand in hand

A first step is to acknowledge that sound ethics and good research practice go hand in hand. The international stem-cell community showed some reluctance to distance itself from Hwang when serious questions were raised in this journal in May 2004 about the manner in which eggs had been obtained for his experiments (see *Nature* **429**, 3; 2004).

As soon as his main US collaborator, Gerald Schatten of the University of Pittsburgh, announced in November that he was bailing out of his collaboration with Hwang (see *Nature* **438**, 262–263; 2005), people began to speculate that Schatten must know there was a problem with the result of the seminal 2004 paper. After all, they inferred, no one would leave a wildly successful research group over ethical transgressions. Or would they?

The leadership of the scientific community has long argued for a solid line to be drawn between ethical transgressions — not informing patients of their rights, sexually harassing staff, coercing junior colleagues, that sort of thing — and actual research misconduct, which refers to the fabrication, falsification or plagiarism of scientific evidence.

In the wake of the Baltimore scandal, a congressionally mandated commission chaired by Kenneth Ryan called for a broader definition

of research misconduct that would embrace some forms of malfeasance beyond fabrication, falsification and plagiarism. His definition didn't cover bioethics, but it did class any breach of research regulations as misconduct.

Ryan's proposal was roundly condemned in the community, which fought a lengthy and successful battle to derail it. Researchers feared that the extension of misconduct investigations to embrace all kinds of professional laxity would lead to endless, fruitless investigation and, in particular, elicit groundless allegations from junior laboratory malcontents.

It is certainly true that there's a distinction between personal misbehaviour in the lab and outright scientific fraud, and it is perhaps

as well that special investigative procedures are retained exclusively for the investigation of the latter. Furthermore, the question of what constitutes an ethical transgression may vary between societies that elect to impose different rules, whereas scientific fraud knows no borders.

In view of the pattern of behaviour that led up to Hwang's disgrace, however, no one should argue ever again that despotism, abuse of junior colleagues, promiscuous authorship on scientific papers or undisclosed payment of research subjects can be tolerated on the grounds of eccentricity or genius. Research ethics matter immensely to the health of the scientific enterprise. Anyone who thinks differently should seek employment in another sphere. ■

Three cheers for peers

Thanks are due to researchers who act as referees, as editors resolve their often contradictory advice.

There is nothing like a high-profile fraud case to encourage journals to reflect on whether the standards and procedures they follow in selecting work for publication are thorough and appropriate. Misconduct creates a negative perception of journals' scientific peer-review processes, and the Hwang fraud saga has already fuelled some misconceptions about how the combination of referees and journal editors actually works.

Peer review remains by far the best available system for scientific quality control, however, and is an ultimately inspiring one at that. *Nature* is hugely grateful for the advice it receives from about 6,000 referees each year — typically two or three referees per paper. Most of their reports contain exactly what we need: a statement of what the referee considers the central message of the paper; an assessment of its significance; and a critique of technical or interpretational weaknesses, either in the work itself or in its presentation.

Between them, these elements add up to a verdict on the work's credibility and robustness. The system, it should be noted, is reliant on trust that what is written in the paper is actually true: it is not designed to detect the tiny minority of papers that are fraudulent.

The lives of editors and editorial boards of all journals are made interesting by the fact that, in many cases, the referees disagree on the verdict. And authors tend to be mightily upset if their papers are rejected when one or more reviewers are positive. Why, they demand to know, should the view of a negative reviewer be allowed to dominate the editors' selection decision?

To shed some light on how these decisions are reached, it is worth reflecting on some case studies of how and why referees differ in their view of papers submitted to *Nature* and the *Nature* research journals.

In one case, an exciting result relied on two techniques and a theoretical interpretation. The theoretical referee was very positive because the work validated an interesting idea. A specialist in one of the techniques was positive because he could find no flaw in its application. But the third referee uncovered a technical shortcoming in the second technique, and the paper was rejected after the editor assessed the significance of the shortcoming.

In another case, one of the referees recommended publication of

a paper, but also pointed out limitations in the value of the finding. The editor concluded that the paper lacked the significance that would justify inclusion in the journal.

On many other occasions, however, the editors' discretion in making a decision results in a paper's publication. In one such case, referees criticized a molecular-biology paper for a lack of mechanistic insight and expressed reservations about the appropriateness of some of the techniques the authors used. But the editors felt that the therapeutic implications of the paper merited publication and, after resolving the technical issues raised by the referees, pushed ahead with publication of what turned out to be a highly cited development.

Another paper concerned the innovative application of chemistry to an environmental problem. But before publication, the editor orchestrated considerable iteration between referees from quite disparate backgrounds to ensure that a common understanding of the paper and its reliability had been established.

In another case, efforts to obtain review of a paper in genetics led to seven refusals to review, one damning review and only one positive review. In this case, the editor identified an experiment that would improve the paper and suggested it, yielding interesting results that were then published and well received.

Only in a minority of cases does every referee agree on whether or not to publish a paper. The above examples illustrate just a few ways in which such differences arise and demonstrate why journals would lose the respect of their authors and readers if they were to act robotically on the referees' advice. Moreover, *Nature* often makes referees aware of what the others are recommending, which can sometimes provide useful feedback to the selection process.

"Only in a minority of cases does every referee agree on whether or not to publish a paper."

We would never claim that *Nature's* decision-making process is perfect. Its imperfections — along with those of every other journal — are among the multifarious reasons why over-reliance on journal publications as a measure of researchers' performance is dangerous. *Nature* has also dished out its fair share of historically embarrassing rejections (see *Nature* 425, 645; 2003). We can only work to ensure that what we publish will do justice to the diversity of expertise that is brought to bear on its selection, and, above all, stand the test of time. ■